Norwegian Centre for
E-health Research

IARIA

Leveraging Machine Learning and Natural Language Processing for Monitoring E-health Publications

Andrius Budrionis, Rune Pedersen, Torbjørn Torsvik, Karianne Lind, Omid Saadatfard

Norwegian Centre for E-health Research

University Hospital of North Norway

Presenter: Andrius Budrionis, Andrius.Budrionis@ehealthresearch.no

About the presenter

Andrius Budrionis, PhD
Senior Researcher, Health data and analytics
Norwegian Centre for E-health Research, University Hospital of North Norway

Background in computer science and software engineering
8+ years experience in e-health

Major focus – clinical data reuse and advanced data analytics. How to deploy and scale machine learning models?

Patient pathways

Citizen services

Health data

Services for health professionals

E-health in Norway

Helsedirektoratet

Direktoratet for e-helse

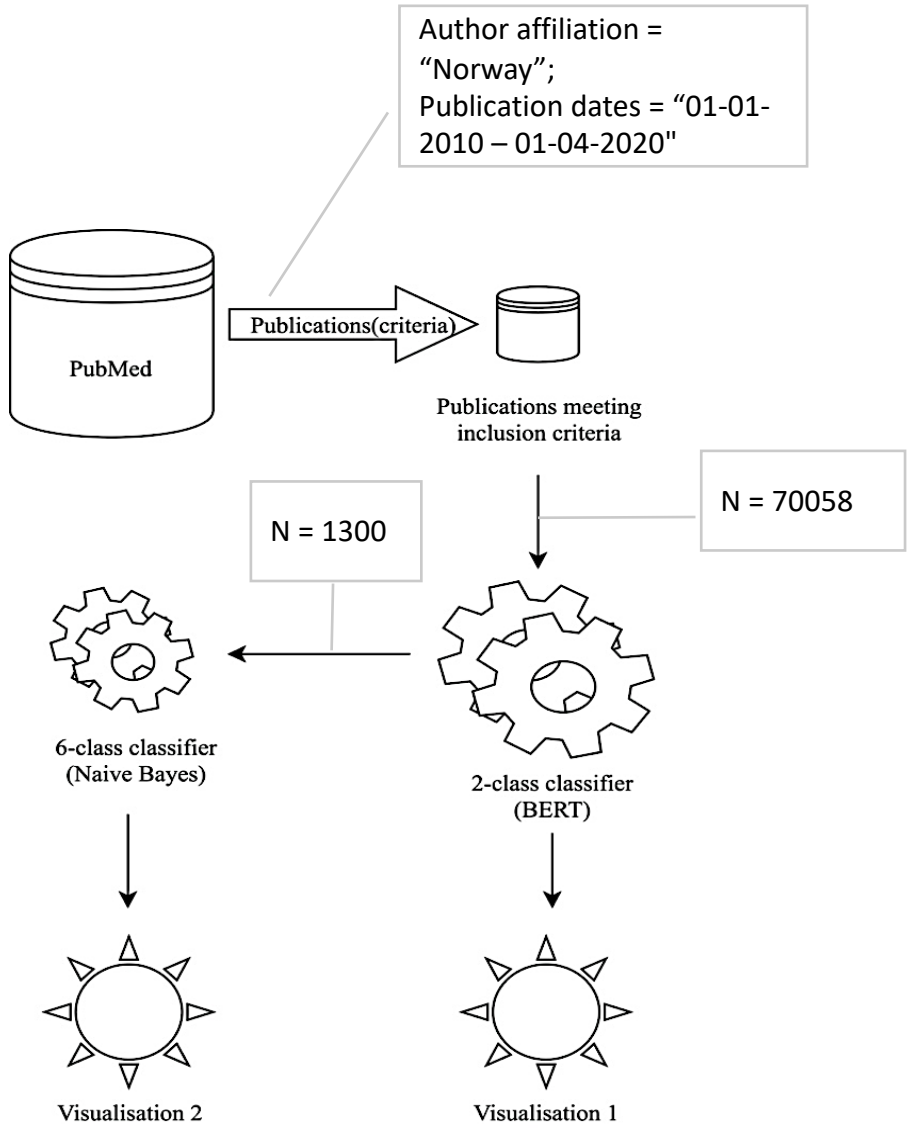
Norwegian Centre for E-health Research

Project plan

1. Digitization of work processes
2. Better continuity of care
3. Better use of health data
4. New ways to provide healthcare
5. Common foundation for digital services
6. National e-health management and increased implementation

Method

Dataset	# e-health	# not-e-health
E-health publication dataset	816	1075
PubMed dataset	25	899



Results. Classification

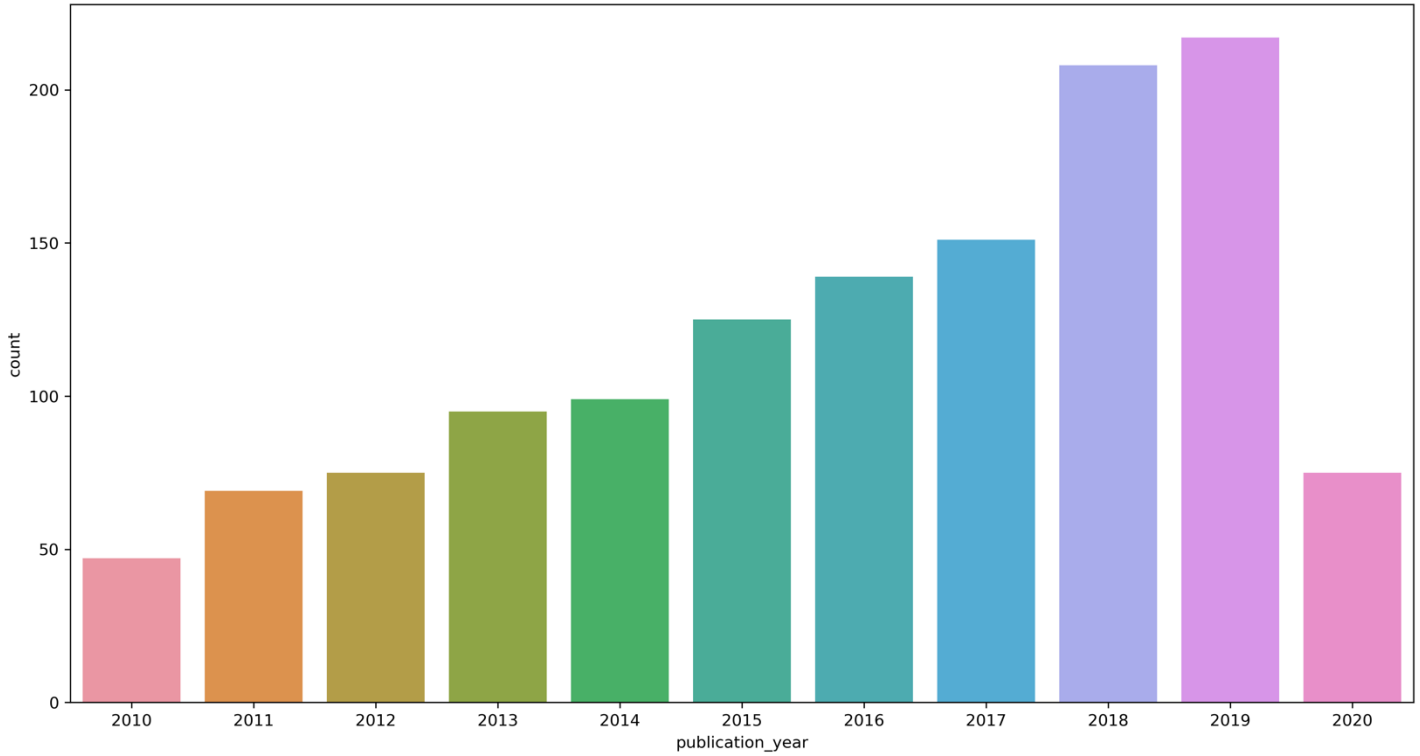
BERT, 2-class classifier

Dataset	Class	Precision	Recall	f-1 score	AUC
E-health publication dataset	Not e-health	0.92	0.88	0.9	0.888
	E-health	0.85	0.90	0.87	
PubMed dataset	Not e-health	0.99	0.99	0.99	0.858
	E-health	0.82	0.72	0.77	

Naive Bayes, 6-class classifier

	Precision	Recall	f1-score
1. Digitization of work processes	0.70	0.58	0.63
2. Better continuity of care	0.61	0.62	0.62
3. Better use of health data	0.62	0.71	0.67
4. New ways to provide healthcare	0.74	0.77	0.75
5. Common foundation for digital services	0.53	0.62	0.57
6. National e-health management and increased implementation	0.66	0.64	0.65

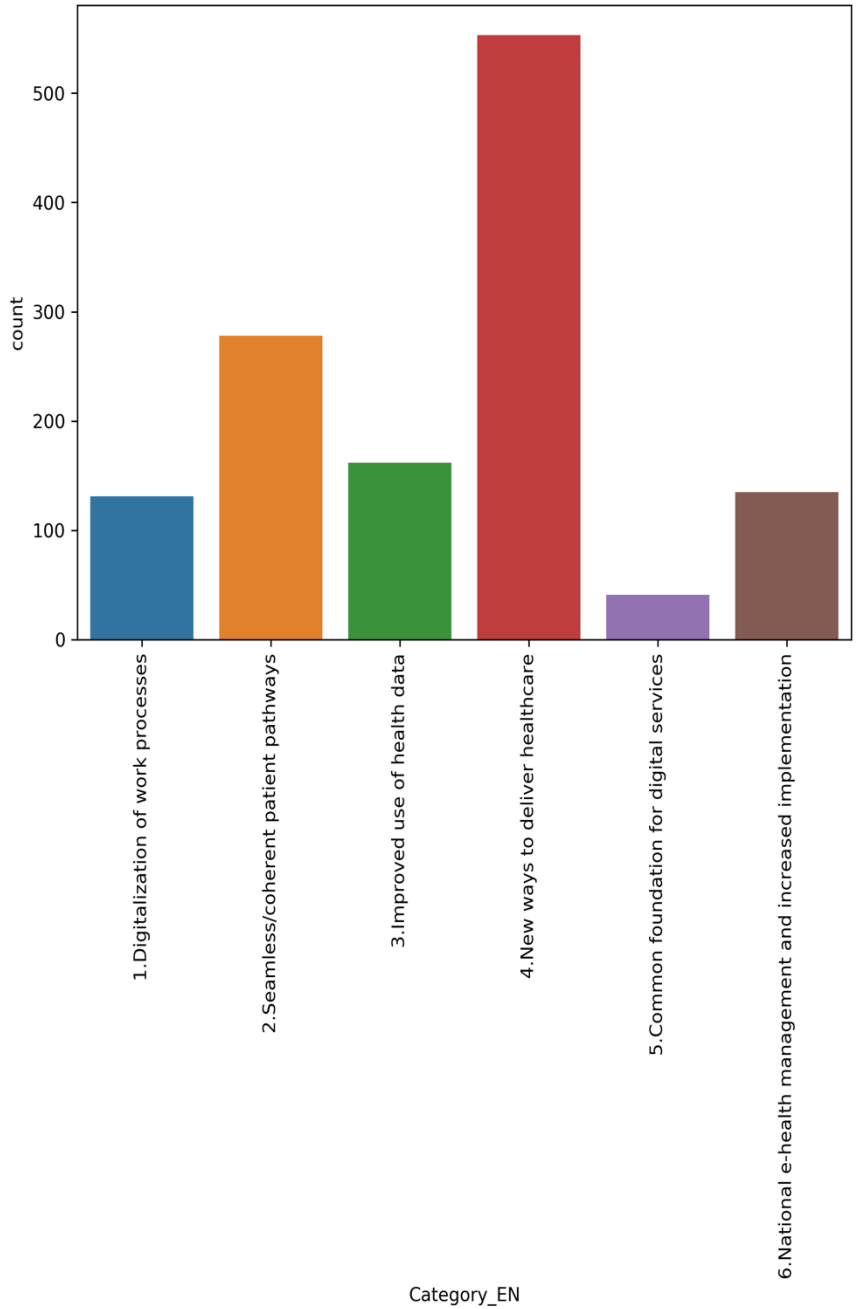
Results. Distribution of e-health publication during last 10 years



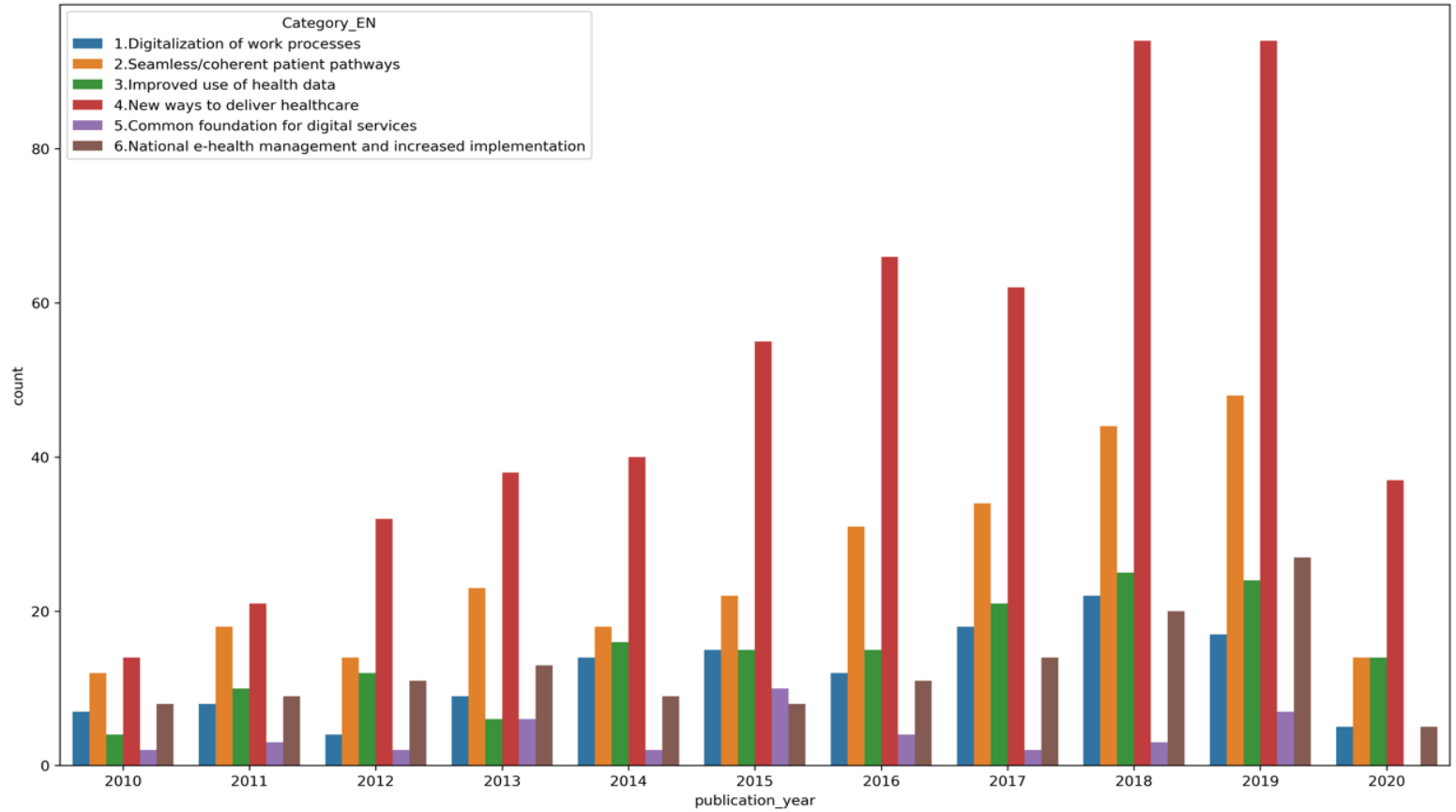
Results. Classification

Top-5 keywords representing each class (sorted by importance):

1. **Digitization of work processes**: nurse, patient, use, care, hospital
2. **Better continuity of care**: patient, care, inform, health, communication
3. **Better use of health data**: data, patient, health, record, predict
4. **New ways to provide healthcare**: diabetes, patient, health, use, social
5. **Common foundation for digital services**: secure, health, standard, information, develop
6. **National e-health management and increased implementation**: telemedicine, health, information, implement, studies



Results. Classified e-health publication stratified yearly



Discussion

- Limitations of data collection
- Classifier performance
- Alternative classification strategies

Way forward

- Integration of other data sources (WoS, Scopus, etc…)
- Use of full-text?
- More granular interactive visualizations

Andrius Budrionis
Norwegian Centre for E-health Research
Andrius.Budrionis@ehealthresearch.no